

Analyse de sensibilité globale appliquée à des modèles dynamiques de systèmes biologiques.

Deux chercheurs de l'unité MaIAGE proposent de nouveaux estimateurs statistiques de trois types d'indices de sensibilité communément utilisés pour l'étude de modèles complexes : indices de Sobol, indices totaux et indices de Shapley. L'approche est illustrée sur deux cas d'études : (i) un réseau Booléen représentant la dynamique d'un réseau de régulation cellulaire ; (ii) un système d'équations différentielles ordinaires modélisant l'infection d'une population bactérienne par un bactériophage. Bien que gourmande en simulations du système, la méthode est bien adaptée pour analyser les modèles mathématiques de systèmes biologiques complexes, pointant notamment les paramètres les plus influents ainsi que leurs potentiels effets combinés sur le comportement dynamique du système.



CONTEXTE ET ENJEUX

L'analyse de sensibilité globale recouvre un ensemble de méthodes mathématiques permettant d'évaluer l'effet d'un ou plusieurs facteurs d'entrée sur une ou plusieurs sorties d'un système complexe. Bien que très utilisés dans certains domaines applicatifs, les indices de sensibilité globale restent majoritairement utilisés en pratique au premier ordre : pour distinguer les facteurs sensibles, donc clés pour l'analyse. Relativement peu explorée, l'estimation d'indices d'ordres supérieurs est intéressante car elle permet de mettre en évidence de possibles effets combinés de certains facteurs d'entrées sur la sortie du système. Ceci devient particulièrement utile lors de l'analyse de systèmes dynamiques déterministes, comme des équations différentielles modélisant des systèmes biologiques complexes : les indices de sensibilité globale deviennent alors un nouvel outil d'analyse pertinent dans l'arsenal du modélisateur.



RESULTATS

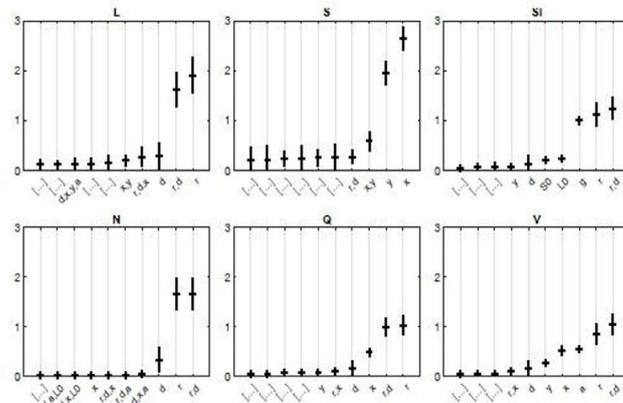
Grâce à une analyse théorique reliant trois types d'indices classiques, nous calculons des estimateurs des indices de Sobol, des indices totaux et des indices de Shapley. Dans tous les cas, nous obtenons directement les matrices de variance-covariance, permettant ainsi d'évaluer précisément la qualité de l'estimation. Nous proposons de plus un ensemble de fonctions informatiques (téléchargeables) permettant de calculer les estimateurs à partir de jeux de données issus de la simulation d'un système quelconque. Nous proposons deux cas d'étude distincts. Le premier est un réseau Booléen modélisant un réseau génétique de décision entre différents destins cellulaires. Les indices permettent alors d'approcher l'ordre de grandeur des vitesses des différents processus biochimiques en jeu, ce qui est souvent inaccessible pour ce type de systèmes. Le second est un système différentiel modélisant l'invasion d'une population bactérienne par un phage. Nous montrons notamment comment les indices calculés peuvent alors fournir des indications précieuses sur la dynamique du système, en révélant par exemple l'interaction clé entre deux paramètres pour décider du type d'équilibre atteint par le système (invasion ou non invasion de la population).



PERSPECTIVES

Ce travail est la collaboration de deux chercheurs en mathématiques appliquées aux spécialités bien différentes : statistique et systèmes dynamiques. Il soulève plusieurs pistes de recherche à l'interface entre ces deux domaines. En particulier, l'utilisation judicieuse d'indices de sensibilité globale, en conjonction avec des techniques d'analyse plus classiques, devrait permettre d'analyser des modèles mathématiques variés, dans différents champs applicatifs d'INRAE.

$$\begin{cases} dL/dt = & r(1 - N/k)L - xL - dL, \\ dS/dt = & r(1 - N/k)S - aVS - dS, \\ dS^l/dt = & r(1 - N/k)S^l - xS^l + gaVS - dS^l, \\ dQ/dt = & x(L + S^l) + (1 - g)aVS - lQ - dQ, \\ dV/dt = & ylQ - aV(N + Q) - dV, \end{cases}$$



En haut : système de cinq équations différentielles modélisant l'invasion d'une population bactérienne S par un phage tempéré V dans le gut de souris. Les variables L et S^l représentent des populations de lysogènes (bactéries infectées avec phage dormant). En bas : indices de Sobol principaux des combinaisons de paramètres sur les 5 variables de l'équilibre du système atteint (N représente la population bactérienne totale $N = S + S^l + L$).
Mazo G. & Tournier L. (2024).

VALORISATION

Nous proposons un ensemble de fonctions informatiques (en Matlab) permettant de calculer nos estimateurs à partir de jeux de données d'un système quelconque. Ce programme, disponible au téléchargement, permet de calculer rapidement les indices pour une quinzaine de paramètres d'entrées. Nos codes sont libres d'accès : leur diffusion dans la communauté permettra d'ouvrir de nouvelles perspectives de collaborations dans d'autres domaines applicatifs.

RÉFÉRENCES

> Mazo G. & Tournier L. (2024) - An inference method for global sensitivity analysis. *Technometrics*, 1-31.
<https://doi.org/10.1080/00401706.2024.2431113>



Contact : Gildas Mazo, Laurent Tournier
Unité : MalAGE
Département : MathNum
Centre INRAE : Île de France Jouy-en-Josas
INRAE 2030 Orientations scientifiques (OS) et Orientations de politique générale (OP) : OS 5.1
Grands Objectifs Scientifiques (GOS) MathNum : GOS 1
Mots-clés : Analyse de sensibilité globale, Indice de Sobol, Equation différentielle ordinaire, Réseau Booléen, Biologie des systèmes